

Canonical dual method for mixed integer fourth-order polynomial minimization problems with fixed cost terms

Zhong Jin · David Y Gao

Abstract we study a canonical duality method to solve a mixed-integer nonconvex fourth-order polynomial minimization problem with fixed cost terms. This constrained nonconvex problem can be transformed into a continuous concave maximization dual problem without duality gap. The global optimality conditions are proposed and the existence and uniqueness criteria are discussed. Application to a decoupled mixed-integer problem is illustrated and analytic solution for a global minimum is obtained under some suitable conditions. Several examples are given to show the method is effective.

Keywords Canonical duality theory · Mixed-integer programming · Fixed cost objective function · Global optimization

Mathematics Subject Classification 90C, 49N

1 Introduction

In this paper, we consider mixed integer minimization problem as follows:

$$(\mathcal{P}_b) : \quad \min\{P(x, v) = Q(x) + W(x) - f^T v \mid (x, v) \in \chi_v\} \quad (1)$$

where $Q(x) = \frac{1}{2}x^T A x - c^T x$, $W(x) = \frac{1}{2}(\frac{1}{2}x^T B x - \alpha)^2$, $A \in \mathbb{R}^{n \times n}$ is a symmetric matrix and $B \in \mathbb{R}^{n \times n}$ is a symmetric positive semi-definite matrix, $\alpha > 0$ is a real number, $c, f \in \mathbb{R}^n$ are two given vectors, $v \in \{0, 1\}^n$ means fixed cost variable, and

$$\chi_v = \{(x, v) \in \mathbb{R}^n \times \{0, 1\}^n \mid -v \leq x \leq v\}. \quad (2)$$

The constraint $x \in [-v, v]$ with $v \in \{0, 1\}^n$ is called as fixed charge constraint [1] and problem \mathcal{P}_b belongs to the class of fixed charge problems, which have attracted considerable attention and extensive study in the integer programming literature in recent years [2–10]. This class of problem has many practical applications in a variety of problems, including facility location [2], network design [3, 5, 7], lot-sizing application [4, 6], and transportation [6, 8–10]. The literature provides only a few exact methods for solving fixed charge problems, such as branch-and-bound type methods, then a good deal of effort has been devoted to finding approximate solutions to fixed charge problems by heuristic methods. Although these problems can be written as mixed integer programs, their computational requirements usually increase exponentially with the size of the problem due to the existence of fixed cost terms in its objective function.

Canonical duality theory, as a breakthrough methodological theory, was originally developed from infinite dimensional nonconvex systems [11]. It has been applied successfully for solving a class of challenging problems in global optimization and nonconvex analysis, such as quadratic problems [12–14], box constrained problems [15], polynomial optimization [15–17],

Z. Jin(✉)

Department of Mathematics, Shanghai Maritime University, Shanghai, 201306, P.R.China
e-mail: asdjz1234@163.com

D.Y. Gao (✉)

Faculty of Science and Technology, Federation University Australia, Victoria 3353, Australia
e-mail: d.gao@federation.edu.au

transportation problems [10], location problems [18], and integer programming problems [12, 19]. Under some certain conditions, these problems can be solved by canonical dual method to obtain global minimums of the primal problems.

Moreover, problem \mathcal{P}_b is related with some problems solved by canonical duality method. Particularly, if the fixed cost term $f^T v = 0$ and the feasible space is defined by $\chi = \{x \in \mathbb{R}^n \mid \mathcal{L}^l \leq x \leq \mathcal{L}^u\}$ where \mathcal{L}^l and \mathcal{L}^u are two given vectors, then \mathcal{P}_b changes to the nonconvex polynomial programming problem with box constraints in [15]. If $f^T v = 0$ and the feasible space is simply defined by $\chi = \{x \in \mathbb{R}^n \mid x_i = 0 \text{ or } 1, i = 1, \dots, n\}$, \mathcal{P}_b converts to the simple 0-1 quadratic programming problem in [12]. If $W(x) = 0$, \mathcal{P}_b becomes the fixed cost quadratic problem that the canonical duality theory is introduced to solve in [10], where the existence and uniqueness of global optimal solutions are proved.

In this paper, we design a canonical duality algorithm to solve a mixed-integer nonconvex fourth-order polynomial minimization problem with fixed cost terms. The presented method is inspired by the method introduced by David Y. Gao for fixed cost quadratic programs [10] and provides a nontrivial extension of his work to polynomial minimization problem. Based on the geometrically admissible operators of [10, 15], we define a new geometrically admissible operator and transform the fourth-order polynomial minimization problem in \mathbb{R}^{2n} into a continuous concave maximization dual problem in \mathbb{R}^{n+1} without duality gap. We give the global optimality conditions and obtain the existence and uniqueness criteria. Furthermore, we discuss the application to decouple problem. Some numerical experiments are given to show our method is effective.

This paper is organised as follows. In the next section, the canonical dual problem for \mathcal{P}_b is formulated, and the corresponding complementary-dual principle is showed. The global optimality criteria are put forward in Section 3, and the existence and uniqueness criteria are proposed in Section 4. We apply our method to decoupled problem in Section 5 and illustrate the effectiveness of our method by some examples in Section 6. Some conclusions and further work are given in the last section.

2 Canonical dual problem

We propose and describe the canonical dual problem of \mathcal{P}_b without duality gap in this section. First, the box constraints $-v \leq x \leq v$, $v \in \{0, 1\}^n$ in the primal problem can be rewritten as relaxed quadratic form:

$$x \circ x \leq v, \quad v \circ (v - e) \leq 0, \quad (3)$$

where $e = \{1\}^n$ is an n -vector of all ones and notation $x \circ v = (x_1 v_1, x_2 v_2, \dots, x_n v_n)$ denotes the Hadamard product between any two vectors $x, v \in \mathbb{R}^n$. So \mathcal{P}_b can be reformulated to the relaxed problem in the following:

$$(\mathcal{P}_r): \quad \min\{P(x, v) = \frac{1}{2}x^T A x - c^T x + \frac{1}{2}\left(\frac{1}{2}x^T B x - \alpha\right)^2 - f^T v \mid x \circ x \leq v, v \circ (v - e) \leq 0\}. \quad (4)$$

In our paper we introduce a so-called geometrically admissible operator

$$y = \Lambda(x, v) = \begin{pmatrix} \xi \\ \epsilon \end{pmatrix} = \begin{pmatrix} \xi \\ \epsilon_1 \\ \epsilon_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}x^T B x - \alpha \\ x \circ x - v \\ v \circ v - v \end{pmatrix} \in \mathbb{R}^{2n+1}, \quad (5)$$

and it is trivial to see that the constraints (3) are equivalent to $\epsilon \leq 0$.

Define

$$V(y) = \frac{1}{2}\xi^2 + \Psi(\epsilon), \quad (6)$$

where

$$\Psi(\epsilon) = \begin{cases} 0 & \text{if } \epsilon \leq 0 \in \mathbb{R}^{2n}, \\ +\infty & \text{otherwise,} \end{cases} \quad (7)$$

and let $U(x, v) = -Q(x) + f^T v$, then unconstrained canonical form of relaxed primal problem (\mathcal{P}_r) can be obtained:

$$(\mathcal{P}_c): \quad \min\{\Pi(x, v) = V(\Lambda(x, v)) - U(x, v) \mid x \in \mathbb{R}^n, v \in \mathbb{R}^n\}. \quad (8)$$

Set $y^* = \begin{pmatrix} \varsigma \\ \sigma \end{pmatrix} = \begin{pmatrix} \varsigma \\ \sigma_1 \\ \sigma_2 \end{pmatrix}$ be the dual vector of $y = \begin{pmatrix} \xi \\ \epsilon \end{pmatrix}$ with $\epsilon \leq 0$, then sup-Fenchel conjugate of $V(y)$ can be defined as

$$\begin{aligned} V^\sharp(y^*) &= \sup_{y \in \mathbb{R}^{1+2n}} \{y^T y^* - V(y)\} \\ &= \sup_{\xi \geq -\alpha} \{\xi \varsigma - \frac{1}{2} \xi^2\} + \sup_{\epsilon_1 \in \mathbb{R}^n} \sup_{\epsilon_2 \in \mathbb{R}^n} \{\epsilon_1^T \sigma_1 + \epsilon_2^T \sigma_2 - \Psi(\epsilon)\} \\ &= \frac{1}{2} \varsigma^2 + \Psi^\sharp(\sigma), \text{ s.t. } \varsigma \geq -\alpha, \end{aligned} \quad (9)$$

where

$$\Psi^\sharp(\sigma) = \begin{cases} 0 & \text{if } \sigma_1 \geq 0 \subseteq \mathbb{R}^n, \sigma_2 \geq 0 \subseteq \mathbb{R}^n \\ +\infty & \text{otherwise.} \end{cases} \quad (10)$$

Now the extended canonical duality relations can be given from the theory of convex analysis:

$$y^* \in \partial V(y) \Leftrightarrow y \in \partial V^\sharp(y^*) \Leftrightarrow V(y) + V^\sharp(y^*) = y^T y^*, \quad (11)$$

and when $\epsilon \leq 0$ and $\sigma \geq 0$ we have

$$\epsilon_1^T \sigma_1 + \epsilon_2^T \sigma_2 = \frac{1}{2} (\xi - \varsigma)^2 \geq 0,$$

then it yields that

$$\epsilon_1^T \sigma_1 = 0 \text{ and } \epsilon_2^T \sigma_2 = 0. \quad (12)$$

Apparently it holds from $\epsilon_2^T \sigma_2 = \sigma_2^T (v \circ v - v) = 0$ in (12) that $v \circ v - v = 0$ for $\forall \sigma_2 > 0$.

By replacing $V(\Lambda(x, v))$ in (8) by the Fenchel-Young equality $V(\Lambda(x, v)) = \Lambda(x, v)^T y^* - V^\sharp(y^*)$, the total complementary function $\Xi(x, v, \varsigma, \sigma)$ corresponding to (\mathcal{P}_c) can be defined in the following:

$$\begin{aligned} \Xi(x, v, \varsigma, \sigma) &= \Lambda(x, v)^T y^* - V^\sharp(y^*) - U(x, v) \\ &= \frac{1}{2} x^T G(\varsigma, \sigma_1) x - c^T x + v^T \text{Diag}(\sigma_2) v - (f + \sigma_1 + \sigma_2)^T v \\ &\quad - \frac{1}{2} \varsigma^2 - \alpha \varsigma - \Psi^\sharp(\sigma), \end{aligned} \quad (13)$$

where

$$G(\varsigma, \sigma_1) = A + \varsigma B + 2 \text{Diag}(\sigma_1), \quad (14)$$

and $\text{Diag}(\sigma_1)$ denotes a diagonal matrix with diagonal elements $(\sigma_1)_i$, $i = 1, \dots, n$. Then we get the canonical dual function:

$$\Pi^d(\varsigma, \sigma) = \text{sta} \{ \Xi(x, v, \varsigma, \sigma) \mid x \in \mathbb{R}^n, v \in \mathbb{R}^n \} = U^\Lambda(\varsigma, \sigma) - V^\sharp(\varsigma, \sigma), \quad (15)$$

where $U^\Lambda(\varsigma, \sigma) = \text{sta} \{ \Lambda(x, v)^T y^* - U(x, v) \mid x \in \mathbb{R}^n, v \in \mathbb{R}^n \}$ is the Λ -conjugate transformation and the notation $\text{sta}\{\cdot\}$ represents calculating stationary point with respect to (x, v) . Correspondingly, introducing a dual feasible space

$$S_\sharp = \{(\varsigma, \sigma) \in \mathbb{R}^1 \times \mathbb{R}^{2n} \mid \varsigma \geq -\alpha, \sigma_1 \geq 0, \sigma_2 > 0, c \in \mathcal{C}_{ol}(G(\varsigma, \sigma_1))\}, \quad (16)$$

where $\mathcal{C}_{ol}(G)$ denotes the column space of G , we can construct the canonical dual function as follows:

$$\begin{aligned} \Pi^d(\varsigma, \sigma) &= U^\Lambda(\varsigma, \sigma) - \frac{1}{2} \varsigma^2 \\ &= -\frac{1}{2} c^T G^+(\varsigma, \sigma_1) c - \frac{1}{4} \sum_{i=1}^n \frac{1}{(\sigma_2)_i} (f_i + (\sigma_1)_i + (\sigma_2)_i)^2 \\ &\quad - \frac{1}{2} \varsigma^2 - \alpha \varsigma, \quad \forall (\varsigma, \sigma) \in S_\sharp, \end{aligned} \quad (17)$$

where G^+ stands for the Moore-Penrose generalized inverse of G . Then it leads to the canonical dual problem for the primal problem (\mathcal{P}_b) as

$$(\mathcal{P}^\sharp) : \quad \max \{ \Pi^d(\zeta, \sigma) \mid (\zeta, \sigma) \in S_\sharp \}. \quad (18)$$

Next we present the complementary-dual principle and for simplicity we denote $t \odot s = \{t_i/s_i\}^n$ for any given n-vectors $t = \{t_i\}^n$ and $s = \{s_i\}^n$.

Theorem 1 *If $(\bar{\zeta}, \bar{\sigma}) \in S_\sharp$ is a KKT point of $\Pi^d(\zeta, \sigma)$, then the vector (\bar{x}, \bar{v}) is feasible to the primal problem (\mathcal{P}_b) and we have*

$$P(\bar{x}, \bar{v}) = \Xi(\bar{x}, \bar{v}, \bar{\zeta}, \bar{\sigma}) = \Pi^d(\bar{\zeta}, \bar{\sigma}), \quad (19)$$

where

$$\bar{x} = \bar{x}(\bar{\zeta}, \bar{\sigma}_1) = G^+(\bar{\zeta}, \bar{\sigma}_1)c, \quad (20)$$

$$\bar{v} = \bar{v}(\bar{\sigma}) = \frac{1}{2}(f + \bar{\sigma}_1 + \bar{\sigma}_2) \odot \bar{\sigma}_2. \quad (21)$$

Proof: Introducing lagrange multiplier $\epsilon = (\epsilon_0, \epsilon_1, \epsilon_2) \in \mathbb{R}^1 \times \mathbb{R}^n \times \mathbb{R}^n$ with the respective three inequalities in (16), we have the lagrangian function Θ for problem (\mathcal{P}^\sharp) :

$$\Theta(\zeta, \sigma, \epsilon_0, \epsilon_1, \epsilon_2) = \Pi^d(\zeta, \sigma) - \epsilon_0(\zeta + \alpha) - \epsilon_1^T \sigma_1 - \epsilon_2^T \sigma_2. \quad (22)$$

It follows from the criticality conditions

$$\nabla_{\sigma_1} \Theta(\bar{\zeta}, \bar{\sigma}, \epsilon_0, \epsilon_1, \epsilon_2) = 0, \quad \nabla_{\sigma_2} \Theta(\bar{\zeta}, \bar{\sigma}, \epsilon_0, \epsilon_1, \epsilon_2) = 0$$

that

$$\epsilon_1 = \nabla_{\sigma_1} \Pi^d(\bar{\zeta}, \bar{\sigma}) = \bar{x}(\bar{\zeta}, \bar{\sigma}_1) \odot \bar{x}(\bar{\zeta}, \bar{\sigma}_1) - \bar{v}(\bar{\sigma}) \quad (23)$$

$$\epsilon_2 = \nabla_{\sigma_2} \Pi^d(\bar{\zeta}, \bar{\sigma}) = \bar{v}(\bar{\sigma}) \odot \bar{v}(\bar{\sigma}) - \bar{v}(\bar{\sigma}), \quad (24)$$

where $\bar{x}(\bar{\zeta}, \bar{\sigma}_1) = G^+(\bar{\zeta}, \bar{\sigma}_1)c$ and $\bar{v}(\bar{\sigma}) = \frac{1}{2}(f + \bar{\sigma}_1 + \bar{\sigma}_2) \odot \bar{\sigma}_2$. Then the corresponding KKT conditions include

$$\bar{\sigma}_1^T \epsilon_1 = 0 \text{ and } \bar{\sigma}_2^T \epsilon_2 = 0, \quad (25)$$

where $\bar{\sigma}_1 \geq 0$, $\bar{\sigma}_2 > 0$, $\epsilon_1 \leq 0$ and $\epsilon_2 \leq 0$. By $\epsilon_1 \leq 0$ and (23), we have $\bar{x} \odot \bar{x} \leq \bar{v}$. Clearly, together with $\bar{\sigma}_2 > 0$ and (24), $\bar{\sigma}_2^T(\bar{v} \odot \bar{v} - \bar{v}) = 0$ in (25) implies that $\bar{v} \odot \bar{v} = \bar{v}$. So when $(\bar{\zeta}, \bar{\sigma})$ is a KKT point of the problem $\Pi^d(\zeta, \sigma)$, (\bar{x}, \bar{v}) is feasible point of (\mathcal{P}_b) .

Furthermore, $\bar{v} = \frac{1}{2}(f + \bar{\sigma}_1 + \bar{\sigma}_2) \odot \bar{\sigma}_2$ implies

$$\sum_{i=1}^n (\bar{\sigma}_2)_i \left(\frac{1}{2} \frac{f_i + (\bar{\sigma}_1)_i + (\bar{\sigma}_2)_i}{(\bar{\sigma}_2)_i} \right)^2 = \bar{\sigma}_2^T(\bar{v} \odot \bar{v}) = \bar{v}^T \text{Diag}(\bar{\sigma}_2) \bar{v}$$

and

$$2\bar{v}^T \text{Diag}(\bar{\sigma}_2) \bar{v} = \bar{v}^T(f + \bar{\sigma}_1 + \bar{\sigma}_2),$$

then with $\bar{x} = G^+(\bar{\zeta}, \bar{\sigma}_1)c$ and $\Psi^\sharp(\bar{\sigma}) = 0$, from (17) we get

$$\begin{aligned} \Pi^d(\bar{\zeta}, \bar{\sigma}) &= \frac{1}{2} \bar{x}^T G(\bar{\zeta}, \bar{\sigma}_1) \bar{x} - c^T \bar{x} - \bar{v}^T \text{Diag}(\bar{\sigma}_2) \bar{v} - \frac{1}{2} \bar{\zeta}^2 - \alpha \bar{\zeta} \\ &= \frac{1}{2} \bar{x}^T G(\bar{\zeta}, \bar{\sigma}_1) \bar{x} - c^T \bar{x} + \bar{v}^T \text{Diag}(\bar{\sigma}_2) \bar{v} - \bar{v}^T(f + \bar{\sigma}_1 + \bar{\sigma}_2) - \frac{1}{2} \bar{\zeta}^2 - \alpha \bar{\zeta} \\ &= \Xi(\bar{x}, \bar{v}, \bar{\zeta}, \bar{\sigma}), \end{aligned} \quad (26)$$

where $(\bar{\zeta}, \bar{\sigma}) \in S_\sharp$. By (14), (26), $\bar{v}^T \text{Diag}(\bar{\sigma}_2) \bar{v} = \bar{\sigma}_2^T(\bar{v} \odot \bar{v})$ and $\bar{x}^T \text{Diag}(\bar{\sigma}_1) \bar{x} = \bar{\sigma}_1^T(\bar{x} \odot \bar{x})$, we have

$$\begin{aligned} \Pi^d(\bar{\zeta}, \bar{\sigma}) &= \Xi(\bar{x}, \bar{v}, \bar{\zeta}, \bar{\sigma}) \\ &= \frac{1}{2} \bar{x}^T \left(A + \bar{\zeta} B + 2 \text{Diag}(\bar{\sigma}_1) \right) \bar{x} - c^T \bar{x} + \bar{\sigma}_2^T(\bar{v} \odot \bar{v}) - \bar{v}^T(f + \bar{\sigma}_1 + \bar{\sigma}_2) \\ &\quad - \frac{1}{2} \bar{\zeta}^2 - \alpha \bar{\zeta} \\ &= \frac{1}{2} \bar{x}^T A \bar{x} - c^T \bar{x} - f^T \bar{v} + \bar{\sigma}_1^T(\bar{x} \odot \bar{x} - \bar{v}) + \bar{\sigma}_2^T(\bar{v} \odot \bar{v} - \bar{v}) \\ &\quad + \frac{1}{2} \bar{\zeta} \bar{x}^T B \bar{x} - \frac{1}{2} \bar{\zeta}^2 - \alpha \bar{\zeta}. \end{aligned} \quad (27)$$

By substituting $\bar{\sigma}_1^T(\bar{x} \circ \bar{x} - \bar{v}) = 0$ and $\bar{\sigma}_2^T(\bar{v} \circ \bar{v} - \bar{v}) = 0$ into (27), we obtain

$$\Pi^d(\bar{\varsigma}, \bar{\sigma}) = \Xi(\bar{x}, \bar{v}, \bar{\varsigma}, \bar{\sigma}) = \frac{1}{2}\bar{x}^T A\bar{x} - c^T \bar{x} - f^T \bar{v} + \frac{1}{2}\bar{\varsigma}\bar{x}^T B\bar{x} - \frac{1}{2}\bar{\varsigma}^2 - \alpha\bar{\varsigma}. \quad (28)$$

From $\nabla_{\varsigma}\Theta(\bar{\varsigma}, \bar{\sigma}, \epsilon_0, \epsilon_1, \epsilon_2) = 0$, together with $\bar{x} = G^+(\bar{\varsigma}, \bar{\sigma}_1)c$ it follows that

$$\frac{1}{2}\bar{x}^T B\bar{x} - \bar{\varsigma} - \alpha - \epsilon_0 = 0, \quad (29)$$

and the accompanying KKT conditions include

$$\bar{\varsigma} + \alpha \geq 0, \quad \epsilon_0 \leq 0, \quad \epsilon_0(\bar{\varsigma} + \alpha) = 0. \quad (30)$$

Suppose $\epsilon_0 < 0$, it holds from $\epsilon_0(\bar{\varsigma} + \alpha) = 0$ that $\bar{\varsigma} + \alpha = 0$, then due to (29) we have

$$\frac{1}{2}\bar{x}^T B\bar{x} - \epsilon_0 = 0,$$

on the other hand, since B is a symmetric positive semi-definite matrix and $\epsilon_0 < 0$, we acquire

$$\frac{1}{2}\bar{x}^T B\bar{x} - \epsilon_0 > 0,$$

which is a contradiction, then $\epsilon_0 = 0$. Thereby from (29), we have

$$\frac{1}{2}\bar{x}^T B\bar{x} - \bar{\varsigma} - \alpha = 0, \quad (31)$$

which implies that $\alpha = \frac{1}{2}\bar{x}^T B\bar{x} - \bar{\varsigma}$ and $\bar{\varsigma} = \frac{1}{2}\bar{x}^T B\bar{x} - \alpha$, then the following equality holds:

$$\frac{1}{2}\bar{\varsigma}^2 + \alpha\bar{\varsigma} = \frac{1}{2}\bar{\varsigma}^2 + \left(\frac{1}{2}\bar{x}^T B\bar{x} - \bar{\varsigma}\right)\bar{\varsigma} = \frac{1}{2}\bar{\varsigma}\bar{x}^T B\bar{x} - \frac{1}{2}\bar{\varsigma}^2 = \frac{1}{2}\bar{\varsigma}\bar{x}^T B\bar{x} - \frac{1}{2}\left(\frac{1}{2}\bar{x}^T B\bar{x} - \alpha\right)^2.$$

Thus it is apparent from (28) that

$$\Pi^d(\bar{\varsigma}, \bar{\sigma}) = \Xi(\bar{x}, \bar{v}, \bar{\varsigma}, \bar{\sigma}) = \frac{1}{2}\bar{x}^T A\bar{x} - c^T \bar{x} + \frac{1}{2}\left(\frac{1}{2}\bar{x}^T B\bar{x} - \alpha\right)^2 - f^T \bar{v} = P(\bar{x}, \bar{v}). \quad (32)$$

The proof is completed. \square

3 Global Optimality Criteria

The global optimality conditions for problem (\mathcal{P}_b) are developed in this section. Firstly, we introduce the following useful feasible space:

$$S_{\sharp}^+ = \{(\varsigma, \sigma) \in \mathbb{R}^1 \times \mathbb{R}^{2n} \mid \varsigma \geq -\alpha, \sigma_1 \geq 0, \sigma_2 > 0, G(\varsigma, \sigma_1) \succ 0\}, \quad (33)$$

where $G(\varsigma, \sigma_1) \succ 0$ means that $G(\varsigma, \sigma_1)$ is a positive definite matrix. It is easy to prove that S_{\sharp}^+ is a convex set. In the following, we use G for short to denote $G(\varsigma, \sigma_1)$.

For convenience, we give the first and second derivatives of function $\Pi^d(\varsigma, \sigma)$:

$$\nabla \Pi^d(\varsigma, \sigma) = \begin{pmatrix} \frac{1}{2}c^T G^{-1} \frac{\partial G}{\partial \varsigma} G^{-1} c - \varsigma - \alpha \\ \left\{ \frac{1}{2}c^T G^{-1} \frac{\partial G}{\partial (\sigma_1)_i} G^{-1} c - \frac{1}{2} \frac{f_i + (\sigma_1)_i + (\sigma_2)_i}{(\sigma_2)_i} \right\}_{i=1}^n \\ \left\{ -\frac{1}{2} \frac{f_j + (\sigma_1)_j + (\sigma_2)_j}{(\sigma_2)_j} + \frac{1}{4} \left(\frac{f_i + (\sigma_1)_i + (\sigma_2)_i}{(\sigma_2)_i} \right)^2 \right\}_{j=1}^n \end{pmatrix}, \quad (34)$$

$$\nabla^2 \Pi^d(\varsigma, \sigma) = -J_1 - J_2 - J_3, \quad (35)$$

in which J_1 , J_2 and $J_3 \in \mathbb{R}^{(2n+1) \times (2n+1)}$ are defined as

$$J_1 = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad J_2 = \begin{bmatrix} Z^T G^{-1} Z & \mathbf{0}_{(n+1) \times n} \\ \mathbf{0}_{n \times (n+1)} & \mathbf{0}_{n \times n} \end{bmatrix}, \quad \text{and } J_3 = \begin{bmatrix} 0 & \mathbf{0}_{1 \times n} & \mathbf{0}_{1 \times n} \\ \mathbf{0}_{n \times 1} & H_{2\sigma_1^2} & H_{\sigma_1\sigma_2} \\ \mathbf{0}_{n \times 1} & H_{\sigma_2\sigma_1} & H_{\sigma_2^2} \end{bmatrix},$$

where

$$\begin{aligned} Z &= \left[\frac{\partial G}{\partial \varsigma} G^{-1} c, \frac{\partial G}{\partial (\sigma_1)_1} G^{-1} c, \frac{\partial G}{\partial (\sigma_1)_2} G^{-1} c, \dots, \frac{\partial G}{\partial (\sigma_1)_n} G^{-1} c \right], \\ H_{2\sigma_1^2} &= \text{Diag} \left\{ -\frac{1}{2(\sigma_2)_i} \right\}, \\ H_{\sigma_1\sigma_2} &= H_{\sigma_2\sigma_1} = \text{Diag} \left\{ \frac{(\sigma_1)_i + f_i}{2(\sigma_2)_i^2} \right\}, \\ H_{\sigma_2^2} &= \text{Diag} \left\{ -\frac{((\sigma_1)_i + f_i)^2}{2(\sigma_2)_i^3} \right\}. \end{aligned}$$

Lemma 2 *The canonical dual function $\Pi^d(\varsigma, \sigma)$ is concave on $S_{\#}^+$.*

Proof: For any given non-zero vector $W = \begin{pmatrix} r \\ s \\ t \end{pmatrix} \in \mathbb{R}^{2n+1}$, where $r \in \mathbb{R}^1$, $s \in \mathbb{R}^n$, $t \in \mathbb{R}^n$,

let $Z_0 = \begin{pmatrix} r \\ s \end{pmatrix}$, by (35), with $G \succ 0$ and $\sigma_2 > 0$ we have

$$\begin{aligned} W^T \nabla^2 \Pi^d(\varsigma, \sigma) W &= -W^T J_1 W - W^T J_2 W - W^T J_3 W \\ &= -r^2 - (ZZ_0)^T G^{-1} (ZZ_0) + \sum_{i=1}^n -\frac{1}{2(\sigma_2)_i} \left(s_i - t_i \frac{(\sigma_1)_i + f_i}{(\sigma_2)_i} \right)^2 \\ &\leq 0, \end{aligned} \tag{36}$$

so the canonical dual function $\Pi^d(\varsigma, \sigma)$ is concave on $S_{\#}^+$. \square

Theorem 3 *Suppose that the vector $\bar{y}^* = (\bar{\varsigma}, \bar{\sigma}) = (\bar{\varsigma}, \bar{\sigma}_1, \bar{\sigma}_2) \in S_{\#}^+$ is a critical point of the dual function $\Pi^d(\varsigma, \sigma)$, then \bar{y}^* is a global maximizer of $\Pi^d(\varsigma, \sigma)$ on $S_{\#}^+$. Let $(\bar{x}, \bar{v}) = (G^{-1}(\bar{\varsigma}, \bar{\sigma}_1)c, \frac{1}{2}(f + \bar{\sigma}_1 + \bar{\sigma}_2) \oslash \bar{\sigma}_2)$, the (\bar{x}, \bar{v}) is a global minimum of $P(x, v)$ on χ_v (i.e., the (\bar{x}, \bar{v}) is a global solution of (\mathcal{P}_b)), and*

$$P(\bar{x}, \bar{v}) = \min_{(x, v) \in \chi_v} P(x, v) = \max_{(\varsigma, \sigma) \in S_{\#}^+} \Pi^d(\varsigma, \sigma) = \Pi^d(\bar{\varsigma}, \bar{\sigma}). \tag{37}$$

Proof: From Lemma 2 the dual function $\Pi^d(\varsigma, \sigma)$ is concave on $S_{\#}^+$, so $\bar{y}^* = (\bar{\varsigma}, \bar{\sigma})$ is a global maximizer of $\Pi^d(\varsigma, \sigma)$ on $S_{\#}^+$, i.e.,

$$\Pi^d(\bar{\varsigma}, \bar{\sigma}) = \max_{(\varsigma, \sigma) \in S_{\#}^+} \Pi^d(\varsigma, \sigma), \tag{38}$$

and this critical point of $\Pi^d(\varsigma, \sigma)$ is a KKT point of $\Pi^d(\varsigma, \sigma)$. Then by Theorem 1, the vector (\bar{x}, \bar{v}) defined by (20) (now $G^+(\bar{\varsigma}, \bar{\sigma}_1) = G^{-1}(\bar{\varsigma}, \bar{\sigma}_1)$) and (21) is a feasible solution to problem (\mathcal{P}_b) and

$$P(\bar{x}, \bar{v}) = \Xi(\bar{x}, \bar{v}, \bar{\varsigma}, \bar{\sigma}) = \Pi^d(\bar{\varsigma}, \bar{\sigma}). \tag{39}$$

As $G \succ 0$ and $\sigma_2 > 0$, we have $\nabla_{(x, v)}^2 \Xi(x, v, \varsigma, \sigma) \succ 0$ and $\nabla_{(\varsigma, \sigma)}^2 \Xi(x, v, \varsigma, \sigma) \preceq 0$. So $\Xi(x, v, \varsigma, \sigma)$ is convex in $(x, v) \in \mathbb{R}^{2n} = \mathbb{R}^n \times \mathbb{R}^n$ and concave in $y^* = (\varsigma, \sigma) \in S_{\#}^+$, then

$$\Pi^d(\varsigma, \sigma) = \min_{(x, v) \in \mathbb{R}^{2n}} \Xi(x, v, \varsigma, \sigma), \tag{40}$$

and $(\bar{x}, \bar{v}, \bar{\varsigma}, \bar{\sigma}) = (\bar{x}, \bar{v}, \bar{\varsigma}, \bar{\sigma})$ is a saddle point of the total complementary function $\Xi(\bar{x}, \bar{v}, \bar{\varsigma}, \bar{\sigma})$ on $\mathbb{R}^{2n} \times S_{\#}^+$, thereby the saddle min-max duality theory holds

$$\max_{(\varsigma, \sigma) \in S_{\#}^+} \min_{(x, v) \in \mathbb{R}^{2n}} \Xi(x, v, \varsigma, \sigma) = \min_{(x, v) \in \mathbb{R}^{2n}} \max_{(\varsigma, \sigma) \in S_{\#}^+} \Xi(x, v, \varsigma, \sigma). \quad (41)$$

Combining (39), (38), (40) and (41), we get

$$P(\bar{x}, \bar{v}) = \min_{(x, v) \in \mathbb{R}^{2n}} \max_{(\varsigma, \sigma) \in S_{\#}^+} \Xi(x, v, \varsigma, \sigma), \quad (42)$$

where $(\bar{x}, \bar{v}) \in \chi_v$.

By (5), $\Lambda(x, v) = \begin{pmatrix} \xi \\ \epsilon \end{pmatrix} = \begin{pmatrix} \xi \\ \epsilon_1 \\ \epsilon_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{2}x^T Bx - \alpha \\ x \circ x - v \\ v \circ v - v \end{pmatrix}$. When $(\varsigma, \sigma) \in S_{\#}^+$, together with (9) and (10), $V^{\#}(y^*) = \frac{1}{2}\varsigma^2$. From (13) and $U(x, v) = -Q(x) + f^T v$, when $(\varsigma, \sigma) \in S_{\#}^+$ we obtain

$$\begin{aligned} \Xi(x, v, \varsigma, \sigma) &= \Lambda(x, v)^T y^* - V^{\#}(y^*) - U(x, v) \\ &= \begin{pmatrix} \xi \\ \epsilon_1 \\ \epsilon_2 \end{pmatrix}^T \begin{pmatrix} \varsigma \\ \sigma_1 \\ \sigma_2 \end{pmatrix} - \frac{1}{2}\varsigma^2 - (-Q(x) + f^T v) \\ &= Q(x) - f^T v - \frac{1}{2}\varsigma^2 + \xi\varsigma + \epsilon_1^T \sigma_1 + \epsilon_2^T \sigma_2, \end{aligned} \quad (43)$$

(a) Consider $\max_{(\varsigma, \sigma) \in S_{\#}^+} (-\frac{1}{2}\varsigma^2 + \xi\varsigma)$.

$$\max_{(\varsigma, \sigma) \in S_{\#}^+} (-\frac{1}{2}\varsigma^2 + \xi\varsigma) = \frac{1}{2}\xi^2 = \frac{1}{2} \left(\frac{1}{2}x^T Bx - \alpha \right)^2 = W(x). \quad (44)$$

(b) Consider $\max_{(\varsigma, \sigma) \in S_{\#}^+} (\epsilon_1^T \sigma_1 + \epsilon_2^T \sigma_2)$.

Noting $\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} = \begin{pmatrix} x \circ x - v \\ v \circ v - v \end{pmatrix}$, by (3) we define the relaxed quadratic form region

$$\bar{\chi}_v = \{(x, v) \in \mathbb{R}^n \times \mathbb{R}^n \mid x \circ x \leq v, v \circ (v - e) \leq 0\}, \quad (45)$$

then we have $\chi_v \subset \bar{\chi}_v$ and

$$\begin{pmatrix} \epsilon_1 \\ \epsilon_2 \end{pmatrix} \leq 0 \iff (x, v) \in \bar{\chi}_v. \quad (46)$$

It follows from $(\varsigma, \sigma) \in S_{\#}^+$ that $\sigma_1 \geq 0$ and $\sigma_2 > 0$, thus with (46) we have

$$\max_{(\varsigma, \sigma) \in S_{\#}^+} (\epsilon_1^T \sigma_1 + \epsilon_2^T \sigma_2) = \begin{cases} 0 & \text{if } (x, v) \in \bar{\chi}_v, \\ +\infty & \text{otherwise (i.e., if } (x, v) \notin \bar{\chi}_v). \end{cases} \quad (47)$$

Taking (43), (44) and (47) into consideration, we find

$$\begin{aligned} \max_{(\varsigma, \sigma) \in S_{\#}^+} \Xi(x, v, \varsigma, \sigma) &= Q(x) - f^T v + \max_{(\varsigma, \sigma) \in S_{\#}^+} (-\frac{1}{2}\varsigma^2 + \xi\varsigma) + \max_{(\varsigma, \sigma) \in S_{\#}^+} (\epsilon_1^T \sigma_1 + \epsilon_2^T \sigma_2) \\ &= \begin{cases} Q(x) + W(x) - f^T v = P(x, v) & \text{if } (x, v) \in \bar{\chi}_v, \\ +\infty & \text{otherwise.} \end{cases} \end{aligned} \quad (48)$$

Then it holds from (42) and (48) that

$$P(\bar{x}, \bar{v}) = \min_{(x, v) \in \mathbb{R}^{2n}} \max_{(\varsigma, \sigma) \in S_{\#}^+} \Xi(x, v, \varsigma, \sigma) = \min_{(x, v) \in \bar{\chi}_v} P(x, v). \quad (49)$$

Because $\chi_v \subset \bar{\chi}_v$, it yields

$$\min_{(x, v) \in \bar{\chi}_v} P(x, v) \leq \min_{(x, v) \in \chi_v} P(x, v),$$

which with (49) lead to

$$P(\bar{x}, \bar{v}) \leq \min_{(x,v) \in \chi_v} P(x, v),$$

which with $(\bar{x}, \bar{v}) \in \chi_v$ implies

$$P(\bar{x}, \bar{v}) = \min_{(x,v) \in \chi_v} P(x, v). \quad (50)$$

Due to the fact that (50), (39) and (38), we get

$$P(\bar{x}, \bar{v}) = \min_{(x,v) \in \chi_v} P(x, v) = \max_{(\varsigma, \sigma) \in S_{\sharp}^+} \Pi^d(\varsigma, \sigma) = \Pi^d(\bar{\varsigma}, \bar{\sigma}).$$

□

Theorem 3 shows that our fourth-order polynomial mixed-integer minimization problem (\mathcal{P}_b) is canonically dual to the concave maximization problem as follows:

$$(\mathcal{P}_{+}^{\sharp}) : \quad \max \left\{ \Pi^d(\varsigma, \sigma) \mid (\varsigma, \sigma) \in S_{\sharp}^+ \right\}. \quad (51)$$

Noted that $\Pi^d(\varsigma, \sigma)$ is a continuous concave function over a convex feasible space S_{\sharp}^+ , if $(\bar{\varsigma}, \bar{\sigma}) \in S_{\sharp}^+$ is a critical point of $\Pi^d(\varsigma, \sigma)$, then it must be a global maximum point of problem $(\mathcal{P}_{+}^{\sharp})$, and $(\bar{x}, \bar{v}) = (G^{-1}(\bar{\varsigma}, \bar{\sigma}_1)c, \frac{1}{2}(f + \bar{\sigma}_1 + \bar{\sigma}_2) \oslash \bar{\sigma}_2)$ should be a global minimum point of problem (\mathcal{P}_b) .

Using $(\sigma_2)_i = |f_i + (\sigma_1)_i|$ is the solution of $\min_{(\sigma_2)_i > 0} \frac{1}{(\sigma_2)_i} (f_i + (\sigma_1)_i + (\sigma_2)_i)^2$, we have

$$\max_{(\sigma_2)_i > 0} -\frac{1}{4} \frac{1}{(\sigma_2)_i} (f_i + (\sigma_1)_i + (\sigma_2)_i)^2 = -(f_i + (\sigma_1)_i)^+. \quad (52)$$

So for a fixed (ς, σ_1) , let

$$\Pi^g(\varsigma, \sigma_1) = \max_{\sigma_2 > 0} \Pi^d(\varsigma, \sigma) = -\frac{1}{2}c^T G^{-1}c - \sum_{i=1}^n (f_i + (\sigma_1)_i)^+ - \frac{1}{2}\varsigma^2 - \alpha\varsigma, \quad (\varsigma, \sigma_1) \in S_{\varsigma\sigma_1}^+, \quad (53)$$

where

$$S_{\varsigma\sigma_1}^+ = \{(\varsigma, \sigma_1) \in \mathbb{R}^1 \times \mathbb{R}^n \mid \varsigma \geq -\alpha, \sigma_1 \geq 0, G(\varsigma, \sigma_1) \succ 0, f_i + (\sigma_1)_i \neq 0, \forall i = 1, \dots, n\}. \quad (54)$$

Then we can write the canonical dual problem $(\mathcal{P}_{+}^{\sharp})$ to a simple form:

$$(\mathcal{P}_{+}^g) : \quad \max \left\{ \Pi^g(\varsigma, \sigma_1) \mid (\varsigma, \sigma_1) \in S_{\varsigma\sigma_1}^+ \right\}. \quad (55)$$

Moreover, set $\delta(t)^+ = \{\delta_i(t_i)^+\}^n \in \mathbb{R}^n$, where

$$\delta_i(t_i)^+ = \begin{cases} 1 & \text{if } t_i > 0, \\ 0 & \text{if } t_i < 0, \end{cases} \quad i = 1, \dots, n. \quad (56)$$

By Theorem 3, it can be easily to get next theorem about analytic solution to primal problem (\mathcal{P}_b) .

Theorem 4 *Given $A \in \mathbb{R}^{n \times n}$, $B \succeq 0 \in \mathbb{R}^{n \times n}$, $c, f \in \mathbb{R}^n$, if $(\bar{\varsigma}, \bar{\sigma}_1) \in S_{\varsigma\sigma_1}^+$ is a critical point of $\Pi^g(\varsigma, \sigma_1)$, then the vector*

$$(\bar{x}, \bar{v}) = (G^{-1}(\bar{\varsigma}, \bar{\sigma}_1)c, \delta(f + \bar{\sigma}_1)^+) \quad (57)$$

is a global minimum point of (\mathcal{P}_b) .

4 Existence and Uniqueness Criteria

In this section we study certain existence and uniqueness conditions for the canonical dual problem to have a critical point. We first let

$$\bar{S}_{\sharp}^+ = \{(\varsigma, \sigma) \in \mathbb{R}^1 \times \mathbb{R}^{2n} \mid \varsigma \geq -\alpha, \sigma_1 \geq 0, \sigma_2 \geq 0, G \succeq 0\}, \quad (58)$$

$$\partial \bar{S}_{\sharp}^+ = \{(\varsigma, \sigma) \in \bar{S}_{\sharp}^+ \mid \det(G) = 0\}. \quad (59)$$

$$\mathfrak{g}_a = \{(\varsigma, \sigma) \in \mathbb{R}^1 \times \mathbb{R}^{2n} \mid \sigma_2 \geq 0, \det(G) = 0\}. \quad (60)$$

Lemma 5 *If $\mathfrak{g}_a \subset \bar{S}_\#^+$, it holds that $\partial \bar{S}_\#^+ = \mathfrak{g}_a$.*

Proof: It is obvious that $\bar{S}_\#^+$ is a closed convex set. From $\mathfrak{g}_a \subset \bar{S}_\#^+$, we have

$$\{(\varsigma, \sigma_1) \in \mathbb{R}^1 \times \mathbb{R}^n \mid \det(G) = 0\} \subset \{(\varsigma, \sigma_1) \in \mathbb{R}^1 \times \mathbb{R}^n \mid \varsigma \geq -\alpha, \sigma_1 \geq 0, G \succeq 0\}$$

which shows

$$\{(\varsigma, \sigma_1) \in \mathbb{R}^1 \times \mathbb{R}^n \mid G \succeq 0\} \subset \{(\varsigma, \sigma_1) \in \mathbb{R}^1 \times \mathbb{R}^n \mid \varsigma \geq -\alpha, \sigma_1 \geq 0\}.$$

Then it holds

$$\bar{S}_\#^+ = \{(\varsigma, \sigma) \in \mathbb{R}^1 \times \mathbb{R}^{2n} \mid \sigma_2 \geq 0, G \succeq 0\},$$

thus together with (59) we have

$$\partial \bar{S}_\#^+ = \{(\varsigma, \sigma) \in \mathbb{R}^1 \times \mathbb{R}^{2n} \mid \sigma_2 \geq 0, \det(G) = 0\} = \mathfrak{g}_a. \quad (61)$$

□

Motivated by the existence and uniqueness criteria given in [14, 20], we have the following theorem about existence and uniqueness criteria.

Theorem 6 *Given $A \in \mathbb{R}^{n \times n}$ and a symmetric positive semi-definite matrix $B \in \mathbb{R}^{n \times n}$, $\alpha > 0$, $c, f \in \mathbb{R}^n$, such that $S_\#^+ \neq \emptyset$ and $\mathfrak{g}_a \subset S_\#^+$. If for any given $(\varsigma_0, \sigma_0) \in \mathfrak{g}_a$ and $(\varsigma, \sigma) \in S_\#^+$,*

$$\lim_{t \rightarrow 0^+} \Pi^d(\varsigma_0 + t\varsigma, \sigma_0 + t\sigma) = -\infty, \quad (62)$$

then the canonical dual problem (\mathcal{P}_+^d) has at least one critical point $(\bar{\varsigma}, \bar{\sigma}_1) \in S_{\varsigma\sigma_1}^+$ and the vector $(\bar{x}, \bar{v}) = (G^{-1}(\bar{\varsigma}, \bar{\sigma}_1)c, \delta(f + \bar{\sigma}_1)^+)$ is a global optimizer of the primal problem (\mathcal{P}_b) . Furthermore, if A, B are two diagonal matrices and $c_i \neq 0, \forall i = 1, \dots, n$, then the vector (\bar{x}, \bar{v}) is a unique global minimum point of (\mathcal{P}_b) .

Proof: It holds from $\mathfrak{g}_a \subset \bar{S}_\#^+$ and Lemma 5 that $\partial \bar{S}_\#^+ = \mathfrak{g}_a$. By the definition, $\bar{S}_\#^+$ is a closed convex set and its interior is $S_\#^+$. If for any given $(\varsigma, \sigma) \in S_\#^+$

$$\begin{aligned} \lim_{t \rightarrow +\infty} \Pi^d(t\varsigma, t\sigma) &= \lim_{t \rightarrow +\infty} \left(-\frac{1}{2} c^T G^{-1}(t\varsigma, t\sigma_1) c - \frac{1}{4} \sum_{i=1}^n \frac{1}{t(\sigma_2)_i} (f_i + t(\sigma_1)_i + t(\sigma_2)_i)^2 \right. \\ &\quad \left. - \frac{1}{2} t^2 \varsigma^2 - \alpha t \varsigma \right) = -\infty, \end{aligned} \quad (63)$$

which with (62) implies that the function $\Pi^d(\varsigma, \sigma)$ is coercive on the open convex set $S_\#^+$. Therefore, (\mathcal{P}_+^d) has at least one critical point $(\bar{\varsigma}, \bar{\sigma}) \in S_\#^+$ and the vector $(\bar{x}, \bar{v}) = (G^{-1}(\bar{\varsigma}, \bar{\sigma}_1)c, \frac{1}{2}(f + \bar{\sigma}_1 + \bar{\sigma}_2) \oslash \bar{\sigma}_2)$ is a global minimum point of problem (\mathcal{P}_b) . So accordingly (\mathcal{P}_+^d) has at least one critical point $(\bar{\varsigma}, \bar{\sigma}_1) \in S_{\varsigma\sigma_1}^+$ and the vector $(\bar{x}, \bar{v}) = (G^{-1}(\bar{\varsigma}, \bar{\sigma}_1)c, \delta(f + \bar{\sigma}_1)^+)$ is a global minimum point of (\mathcal{P}_b) .

Next we will prove the canonical dual function $\Pi^d(\varsigma, \sigma)$ is strict concave on $S_\#^+$.

From (36), we have

$$W^T \nabla^2 \Pi^d(\varsigma, \sigma) W = -r^2 - (ZZ_0)^T G^{-1}(ZZ_0) + \sum_{i=1}^n -\frac{1}{2(\sigma_2)_i} \left(s_i - t_i \frac{(\sigma_1)_i + f_i}{(\sigma_2)_i} \right)^2. \quad (64)$$

(a) If $Z_0 = \begin{pmatrix} r \\ s \end{pmatrix} = 0$, $W \neq 0$ leads to $t \neq 0$, then with $(\sigma_2)_i > 0$ and $f_i + (\sigma_1)_i \neq 0$, we get

$$W^T \nabla^2 \Pi^d(\varsigma, \sigma) W = \sum_{i=1}^n -\frac{1}{2(\sigma_2)_i} \left(0 - t_i \frac{(\sigma_1)_i + f_i}{(\sigma_2)_i} \right)^2 = \sum_{i=1}^n -\frac{t_i^2}{2((\sigma_2)_i)^3} (f_i + (\sigma_1)_i)^2 < 0.$$

- (b) If $Z_0 = \begin{pmatrix} r \\ s \end{pmatrix} \neq 0$, there are two cases.
- (b.1) If $r \neq 0$, then $W^T \nabla^2 \Pi^d(\varsigma, \sigma) W \leq -r^2 < 0$.
- (b.2) If $r = 0$, then $s \neq 0$ and $Z_0 = \begin{pmatrix} 0 \\ s \end{pmatrix}$, thus

$$-(ZZ_0)^T G^{-1}(ZZ_0) = -4s^T \text{Diag}(x(\varsigma, \sigma_1))^T G^{-1} \text{Diag}(x(\varsigma, \sigma_1))s. \quad (65)$$

Since A and B are two diagonal matrices, then from (14) we know G is a diagonal matrix, which with $G \succ 0$ shows that G^{-1} is a diagonal matrix whose all diagonal elements are positive. Then it follows from $x(\varsigma, \sigma_1) = G^{-1}c$ and $c_i \neq 0, \forall i = 1, \dots, n$ that

$$x(\varsigma, \sigma_1)_i \neq 0, \forall i = 1, \dots, n.$$

By $G^{-1} \succ 0$, we find $\text{Diag}(x(\varsigma, \sigma_1))^T G^{-1} \text{Diag}(x(\varsigma, \sigma_1)) \succ 0$, then it is observed from (65) that

$$-(ZZ_0)^T G^{-1}(ZZ_0) < 0,$$

hence

$$W^T \nabla^2 \Pi^d(\varsigma, \sigma) W \leq -(ZZ_0)^T G^{-1}(ZZ_0) < 0.$$

From above two cases, we can see $\Pi^d(\varsigma, \sigma)$ is strictly concave on $S_{\#}^+$. Thereby $(\mathcal{P}_{+}^{\#})$ has a unique critical point in $S_{\#}^+$, which implies (\mathcal{P}_{+}^g) has a unique critical point in $S_{\varsigma\sigma_1}^+$ and the primal problem has a unique global minimum. \square

5 Application to Decoupled Problem

In this section, we discuss the application of presented theory to the decoupled problems. Consider the decoupled minimization problem as follows:

$$\min \left\{ P(x, v) = \frac{1}{2} \left(\frac{1}{2} \sum_{i=1}^n b_i x_i^2 - \alpha \right)^2 + \sum_{i=1}^n \left(\frac{1}{2} a_i x_i^2 - c_i x_i - f_i v_i \right) \right\} \quad (66)$$

$$s.t. \quad -v_i \leq x_i \leq v_i, v_i \in \{0, 1\}, i = 1, \dots, n. \quad (67)$$

For simplicity, we may take $A = \text{Diag}(a)$ and $B = \text{Diag}(b)$ stand for two diagonal matrices with diagonal elements $a = \{a_i\} \in \mathbb{R}^n$ and $b = \{b_i\} \in \mathbb{R}^n$ respectively. The canonical dual function of above problem has the following simple form

$$\Pi^d(\varsigma, \sigma) = -\frac{1}{2} \sum_{i=1}^n \left(\frac{c_i^2}{a_i + \varsigma b_i + 2(\sigma_1)_i} + \frac{(f_i + (\sigma_1)_i + (\sigma_2)_i)^2}{2(\sigma_2)_i} \right) - \frac{1}{2} \varsigma^2 - \alpha \varsigma. \quad (68)$$

Due to $\nabla \Pi^d(\varsigma, \sigma) = 0$, then when $(\sigma_2)_i \neq 0$ and $c_i \neq 0$, we can obtain the critical points of $\Pi^d(\varsigma, \sigma)$ in the following

$$\varsigma = \frac{1}{2} \sum_{i=1}^n b_i - \alpha, \quad (\sigma_1)_i \in M_i = \left\{ -\frac{1}{2} \left(a_i + \left(\frac{1}{2} \sum_{i=1}^n b_i - \alpha \right) b_i \pm c_i \right) \right\}, \quad (69)$$

$$(\sigma_2)_i \in N_i = \left\{ f_i - \frac{1}{2} \left(a_i + \left(\frac{1}{2} \sum_{i=1}^n b_i - \alpha \right) b_i \pm c_i \right) \right\}, \quad \forall i = 1, \dots, n. \quad (70)$$

For $(\sigma_2)_i > 0$, using Theorem 1, we can derive the accompanying primal solution

$$(x_i, v_i) = \left(-\frac{c_i}{a_i + (\frac{1}{2} \sum_{i=1}^n b_i - \alpha) b_i + 2(\sigma_1)_i}, \frac{f_i + (\sigma_1)_i + (\sigma_2)_i}{2(\sigma_2)_i} \right), \quad \forall i = 1, \dots, n. \quad (71)$$

According to the fact that there are two possible solutions for each component of $\sigma = (\sigma_1, \sigma_2) \in \mathbb{R}^{2n}$, together with (69) and (70), it follows that the canonical dual function Π^d has 2^n critical points. Then from Theorem 3, it is easy to show the global minimum of the primal problem will be arrived at by next theorem.

Theorem 7 Given $A = \text{Diag}(a) \in \mathbb{R}^{n \times n}$ and a symmetric positive semi-definite matrix $B = \text{Diag}(b) \in \mathbb{R}^{n \times n}$, $\alpha > 0$, $c, f \in \mathbb{R}^n$, $c_i \neq 0$ for $\forall i$, if

$$\max M_i > 0 \quad \text{and} \quad \max N_i > 0, \quad \forall i = 1, \dots, n. \quad (72)$$

then $\Pi^d(\zeta, \sigma)$ has a unique critical point

$$(\bar{\zeta}, \bar{\sigma}) = (\bar{\zeta}, \bar{\sigma}_1, \bar{\sigma}_2) = \left(\frac{1}{2} \sum_{i=1}^n b_i - \alpha, \{ \max M_i, i = 1, \dots, n \}, \{ \max N_i, i = 1, \dots, n \} \right) \in S_{\sharp}^+,$$

which is a global maximizer of $\Pi^d(\zeta, \sigma)$ on S_{\sharp}^+ , and

$$(\bar{x}, \bar{v}) = \left(\left\{ \frac{c_i}{|c_i|} \right\}, e \right)$$

is a global minimum of $P(x, v)$ on χ_v .

6 Examples

Now we give a summary of numerical experiments to illustrate our method, where diagonal matrices A and B , vectors f and c are chosen and taken at random. These examples are classified into three cases and in every case we give several representatives. In the first case we consider the decoupled problems satisfying the conditions of Theorem 7, whose results are consistent with Theorem 7 and show our method is promising for decoupled problems. In the second case some general decoupled problems not satisfying Theorem 7 are computed by our method and the global solution are also obtained. In the last case, our method is tested for some general problems. All of performed examples show our method is very effective.

6.1 Case 1: Decoupled Problems satisfying Theorem 7

In the following three decoupled examples, we can verify the conditions of Theorem 7 are satisfied, so a unique critical point $(\bar{\zeta}, \bar{\sigma})$ of $\Pi^d(\zeta, \sigma)$ on S_{\sharp}^+ and a global minimum point (\bar{x}, \bar{v}) of $P(x, v)$ on χ_v are obtained by Theorem 7. For simplicity we denote $\lambda_{\min}(\bar{\zeta}, \bar{\sigma}_1)$ be the smallest eigenvalue of G .

Example 1 Set $\alpha = 10$, $A = \text{Diag}(1, -1, 1, 5, 2)$, $B = \text{Diag}(2, 4, 1, 4, 2)$, $f = (20, 12, -1, 1, 13)$ and $c = (-8, -9, 10, 9, -5)$, in which $n = 5$.

Example 2 Set $\alpha = 20$, $A = \text{Diag}(7, 9, 6, -5, 4, 10, 9, 8)$, $B = \text{Diag}(3, 5, 4, 3, 1, 7, 5, 7)$, $f = (13, -3, 3, 11, 10, 16, 16, 14)$ and $c = (7, -6, 8, -1, -5, 8, -8, 7)$, in which $n = 8$.

Example 3 Set $\alpha = 25$, $A = \text{Diag}(2, 8, 7, 3, 6, 14, 10, 1, -6, 9)$, $B = \text{Diag}(9, 1, 2, 1, 6, 8, 5, 3, 9, 6)$, $f = (6, 1, 4, 13, 6, 15, 17, 20, 3, 16)$ and $c = (19, 14, -9, -9, -8, 17, -22, -14, -8, 18)$, in which $n = 10$.

The values of $\frac{1}{2} \sum_{i=1}^n b_i - \alpha$, M_i and N_i from (69) and (70) are first needed to be computed and are listed in Table 1 as follows. Then it can be found that conditions of Theorem 7 are all satisfied.

| Experiments | $\frac{1}{2} \sum_{i=1}^n b_i - \alpha$ | M_i | $Max M_i$ | N_i | $Max N_i$ |
|-------------|---|---------------------------|-----------|--------------------------|-----------|
| Example 1 | -3.5 | $M_1 = \{7, -1\}$ | 7 | $N_1 = \{27, 19\}$ | 27 |
| | | $M_2 = \{12, 3\}$ | 12 | $N_2 = \{24, 15\}$ | 24 |
| | | $M_3 = \{-3.75, 6.25\}$ | 6.25 | $N_3 = \{-4.75, 5.25\}$ | 5.25 |
| | | $M_4 = \{0, 9\}$ | 9 | $N_4 = \{1, 10\}$ | 10 |
| | | $M_5 = \{5, 0\}$ | 5 | $N_5 = \{18, 13\}$ | 18 |
| Example 2 | -2.5 | $M_1 = \{-3.25, 3.75\}$ | 3.75 | $N_1 = \{9.75, 16.75\}$ | 16.75 |
| | | $M_2 = \{4.75, -1.25\}$ | 4.75 | $N_2 = \{1.75, -4.25\}$ | 1.75 |
| | | $M_3 = \{-2, 6\}$ | 6 | $N_3 = \{1, 9\}$ | 9 |
| | | $M_4 = \{6.75, 5.75\}$ | 6.75 | $N_4 = \{17.75, 16.75\}$ | 17.75 |
| | | $M_5 = \{1.75, -3.25\}$ | 1.75 | $N_5 = \{11.75, 6.75\}$ | 11.75 |
| | | $M_6 = \{-0.25, 7.75\}$ | 7.75 | $N_6 = \{15.75, 23.75\}$ | 23.75 |
| | | $M_7 = \{5.75, -2.25\}$ | 5.75 | $N_7 = \{21.75, 13.75\}$ | 21.75 |
| | | $M_8 = \{1.25, 8.25\}$ | 8.25 | $N_8 = \{15.25, 22.25\}$ | 22.25 |
| Example 3 | 0 | $M_1 = \{-10.5, 8.5\}$ | 8.5 | $N_1 = \{-4.5, 14.5\}$ | 14.5 |
| | | $M_2 = \{-11, 3\}$ | 3 | $N_2 = \{-10, 4\}$ | 4 |
| | | $M_3 = \{1, -8\}$ | 1 | $N_3 = \{5, -4\}$ | 5 |
| | | $M_4 = \{3, -6\}$ | 3 | $N_4 = \{16, 7\}$ | 16 |
| | | $M_5 = \{1, -7\}$ | 1 | $N_5 = \{7, -1\}$ | 7 |
| | | $M_6 = \{-15.5, 1.5\}$ | 1.5 | $N_6 = \{-0.5, 16.5\}$ | 16.5 |
| | | $M_7 = \{6, -16\}$ | 6 | $N_7 = \{23, 1\}$ | 23 |
| | | $M_8 = \{6.5, -7.5\}$ | 6.5 | $N_8 = \{26.5, 12.5\}$ | 26.5 |
| | | $M_9 = \{7, -1\}$ | 7 | $N_9 = \{10, 2\}$ | 10 |
| | | $M_{10} = \{-13.5, 4.5\}$ | 4.5 | $N_{10} = \{2.5, 20.5\}$ | 20.5 |

Table 1: Conditions of Theorem 7 are satisfied in Examples 1-3.

On the one hand, by Theorem 7 we know $\varsigma = \frac{1}{2} \sum_{i=1}^n b_i - \alpha$, $(\bar{\sigma}_1)_i = \max M_i$, $(\bar{\sigma}_2)_i = \max N_i$ and $(\bar{x}, \bar{v}) = \left(\left\{ \frac{c_i}{|c_i|} \right\}, e \right)$, so from Table 1 we can easily get the corresponding $(\bar{\varsigma}, \bar{\sigma})$ and (\bar{x}, \bar{v}) for Examples 1-3 listed in Table 2.

| Experiments | $\bar{\varsigma}$ | $\bar{\sigma}$ | Primal solution (\bar{x}, \bar{v}) |
|-------------|--------------------------|--|--|
| Example 1 | $\bar{\varsigma} = -3.5$ | $\bar{\sigma}_1 = (7, 12, 6.25, 9, 5),$ $\bar{\sigma}_2 = (27, 24, 5.25, 10, 18).$ | $\bar{x} = (-1, -1, 1, 1, -1),$ $\bar{v} = (1, 1, 1, 1, 1).$ |
| Example 2 | $\bar{\varsigma} = -2.5$ | $\bar{\sigma}_1 = (3.75, 4.75, 6, 6.75, 1.75, 7.75, 5.75, 8.25),$ $\bar{\sigma}_2 = (16.75, 1.75, 9, 17.75, 11.75, 23.75, 21.75, 22.25).$ | $\bar{x} = (1, -1, 1, -1, -1, 1, -1, 1)$ $\bar{v} = (1, 1, 1, 1, 1, 1, 1, 1).$ |
| Example 3 | $\bar{\varsigma} = 0$ | $\bar{\sigma}_1 = (8.5, 3, 1, 3, 1, 1.5, 6, 6.5, 7, 4.5),$ $\bar{\sigma}_2 = (14.5, 4, 5, 16, 7, 16.5, 23, 26.5, 10, 20.5).$ | $\bar{x} = (1, 1, -1, -1, -1, 1, -1, -1, -1, 1),$ $\bar{v} = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1).$ |

Table 2: Results of Theorem 7 in Examples 1-3.

On the other hand, we can also get $(\bar{\varsigma}, \bar{\sigma})$ by solving the canonical dual problem (\mathcal{P}_+^\sharp) , then (\bar{x}, \bar{v}) can be computed by $(\bar{x}, \bar{v}) = (G^{-1}(\bar{\varsigma}, \bar{\sigma}_1)c, \frac{1}{2}(f + \bar{\sigma}_1 + \bar{\sigma}_2) \odot \bar{\sigma}_2)$. The corresponding results are listed below:

| Experiments | Dual solution $(\bar{\varsigma}, \bar{\sigma})$ | $H^d(\bar{\varsigma}, \bar{\sigma})$ | $\lambda_{min}(\bar{\varsigma}, \bar{\sigma}_1)$ | Primal solution (\bar{x}, \bar{v}) | $P(\bar{x}, \bar{v})$ |
|-------------|---|--------------------------------------|--|--|-----------------------|
| Example 1 | $\bar{\varsigma} = -3.5,$ $\bar{\sigma}_1 = (7, 12, 6.25, 9, 5),$ $\bar{\sigma}_2 = (27, 24, 5.25, 10, 18).$ | -75.875 | 5 | $\bar{x} = (-1, -1, 1, 1, -1),$ $\bar{v} = (1, 1, 1, 1, 1).$ | -75.875 |
| Example 2 | $\bar{\varsigma} = -2.5,$ $\bar{\sigma}_1 = (3.75, 4.75, 6, 6.75, 1.75, 7.75, 5.75, 8.25),$ $\bar{\sigma}_2 = (16.75, 1.75, 9, 17.75, 11.75, 23.75, 21.75, 22.25).$ | -102.875 | 1 | $\bar{x} = (1, -1, 1, -1, -1, 1, -1, 1),$ $\bar{v} = (1, 1, 1, 1, 1, 1, 1, 1).$ | -102.875 |
| Example 3 | $\bar{\varsigma} = 0,$ $\bar{\sigma}_1 = (8.5, 3, 1, 3, 1, 1.5, 6, 6.5, 7, 4.5),$ $\bar{\sigma}_2 = (14.5, 4, 5, 16, 7, 16.5, 23, 26.5, 10, 20.5).$ | -212 | 8 | $\bar{x} = (1, 1, -1, -1, -1, 1, -1, -1, -1, 1),$ $\bar{v} = (1, 1, 1, 1, 1, 1, 1, 1, 1, 1).$ | -212 |

Table 3: Results by solving problem (\mathcal{P}_+^\sharp) in Examples 1-3.

From Tables 2 and 3, it can be seen that the results from Theorem 7 are consistent with the ones by our canonical dual method. And the fact that $\bar{\varsigma} \geq -\alpha, \bar{\sigma}_1 \geq 0, \bar{\sigma}_2 \geq 0, G(\bar{\varsigma}, \bar{\sigma}_1) \succeq 0$ in every example indicates $(\bar{\varsigma}, \bar{\sigma}) \in S_+^\dagger$. By Theorem 7, the solution of primal problem is obtained. It is verified that our method is promising for decoupled problems when the conditions of Theorem 7 are satisfied for the decoupled problems.

6.2 Part 2: Decoupled Problems where Theorem 7 are not satisfied

It needs to say that there are many decoupled problems not to satisfy Theorem 7. Here we choose three of them to indicate the details. Let $n = 5$ and $\alpha = 10$, diagonal matrices A and B , vectors f and c are chosen and taken at random.

Example 4 Set $A = \text{Diag}(6, 3, 9, 9, 2)$, $B = \text{Diag}(2, 4, 5, 4, 3)$, $f = (5, 4, 4, 20, 9)$ and $c = (1, -9, -6, 3, -5)$.

Example 5 Set $A = \text{Diag}(1, -1, 1, 4, 4)$, $B = \text{Diag}(1, 1, 1, 4, 5)$, $f = (1, -51, -1, -11, -61)$ and $c = (3, 0, 1, -2, 0)$.

Example 6 Set $A = \text{Diag}(5, -1, 2, 5, 1)$, $B = \text{Diag}(5, 2, 2, 1, 4)$, $f = (3, -35, -1, 11, 15)$ and $c = (7, 0, 4, -6, 10)$.

| Experiments | $\frac{1}{2} \sum_{i=1}^n b_i - \alpha$ | M_i | $\text{Max } M_i$ | N_i | $\text{Max } N_i$ |
|-------------|---|------------------------|-------------------|----------------------|-------------------|
| Example 4 | -1 | $M_1 = \{-2.5, -1.5\}$ | -1.5 | $N_1 = \{2.5, 3.5\}$ | 3.5 |
| | | $M_2 = \{5, -4\}$ | 5 | $N_2 = \{9, 0\}$ | 9 |
| | | $M_3 = \{1, -5\}$ | 1 | $N_3 = \{5, -1\}$ | 5 |
| | | $M_4 = \{-4, -1\}$ | -1 | $N_4 = \{16, 19\}$ | 19 |
| | | $M_5 = \{3, -2\}$ | 3 | $N_5 = \{12, 7\}$ | 12 |

Table 4: Conditions of Theorem 7 are not all satisfied in Examples 4.

From Table 4, we find that $\text{Max } M_1 = -1.5 < 0$ which makes the conditions of Theorem 7 are not satisfied. We solve the simple form problem (\mathcal{P}_+^g) instead of the canonical dual problem $(\mathcal{P}_+^\#)$ and use Theorem 4 to obtain the analytic solution (x, v) to primal problem (\mathcal{P}_b) . The corresponding results are listed below:

| Experiments | Dual solution $(\bar{\varsigma}, \bar{\sigma}_1)$ | $\Pi^g(\bar{\varsigma}, \bar{\sigma}_1)$ | $\lambda_{\min}(\bar{\varsigma}, \bar{\sigma}_1)$ | Primal solution (\bar{x}, \bar{v}) | $P(\bar{x}, \bar{v})$ |
|-------------|---|--|---|---|-----------------------|
| Example 4 | $\bar{\varsigma} = -1.82$, $\bar{\sigma}_1 = (0, 6.641, 3.051, 0.641, 4.231)$. | -51.7281 | 2.3593 | $\bar{x} = (0.424, -1, -1, 1, -1)$, $\bar{v} = (1, 1, 1, 1, 1)$. | -51.7281 |

Table 5: Results by solving problem (\mathcal{P}_+^g) in Examples 4.

It is obvious that there exists some $c_i = 0$ in Examples 5 and 6, which does not satisfy the conditions of Theorem 7. We also solve them by the simple form problem (\mathcal{P}_+^g) , whose results are illustrated in Table 6.

| Experiments | Dual solution $(\bar{\varsigma}, \bar{\sigma}_1)$ | $\Pi^g(\bar{\varsigma}, \bar{\sigma}_1)$ | $\lambda_{\min}(\bar{\varsigma}, \bar{\sigma}_1)$ | Primal solution (\bar{x}, \bar{v}) | $P(\bar{x}, \bar{v})$ |
|-------------|---|--|---|---|-----------------------|
| Example 5 | $\bar{\varsigma} = -7$, $\bar{\sigma}_1 = (4.5, 34.987, 3.5, 13, 54.367)$. | 32.5 | 1 | $\bar{x} = (1, 0, 1, -1, 0)$, $\bar{v} = (1, 0, 1, 1, 0)$. | 32.5 |
| Example 6 | $\bar{\varsigma} = -4$, $\bar{\sigma}_1 = (11, 5.612, 5, 2.5, 12.5)$. | -40.5 | 2.22424 | $\bar{x} = (1, 0, 1, -1, 1)$, $\bar{v} = (1, 0, 1, 1, 1)$. | -40.5 |

Table 6: Results by solving problem (\mathcal{P}_+^g) in Examples 5 and 6.

From Tables 5-6, we have $\bar{\varsigma} \geq -\alpha$, $\bar{\sigma}_1 \geq 0$, $G(\bar{\varsigma}, \bar{\sigma}_1) \succ 0$, $f_i + (\bar{\sigma}_1)_i \neq 0, \forall i = 1, \dots, n$ in every example, so $(\bar{\varsigma}, \bar{\sigma}_1) \in S_{\bar{\varsigma}\bar{\sigma}_1}^+$. By Theorem 4, the solution of primal problem is obtained. Thus when the conditions of Theorem 7 are not all satisfied, our method is also effective for the decoupled problem.

6.3 Part 3: General Nonconvex Problems

For general nonconvex problems in this part, we solve the simple form problem (\mathcal{P}_+^g) and use Theorem 4 to obtain the analytic solution (x, v) to primal problem (\mathcal{P}_b) . Two general nonconvex examples not decoupled are tested and the corresponding results are listed below:

Example 7 Set

$$A = \begin{bmatrix} 4 & 0 & 1 \\ 0 & -4 & -6 \\ 1 & -6 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 7 & -3 & -4 \\ -3 & 8 & 2 \\ -4 & 2 & 10 \end{bmatrix},$$

$\alpha = 8$, $f = (3, 2, 3)$ and $c = (10, 6, 7)$.

Example 8 Set

$$A = \begin{bmatrix} 15 & 3 & -3 & -2 & -4 \\ 3 & 21 & -5 & 0 & 2 \\ -3 & -5 & 12 & 0 & 2 \\ -2 & 0 & 0 & 14 & 3 \\ -4 & -2 & 2 & 3 & 6 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 13 & 2 & -4 & 4 & -6 \\ 2 & 6 & -4 & 1 & -2 \\ -4 & -4 & 6 & 0 & -3 \\ 4 & 1 & 0 & 7 & -7 \\ -6 & -2 & -3 & -7 & 21 \end{bmatrix},$$

$\alpha = 4$, $f = (6, -2, 5, 4, 10)$ and $c = (7, -3, 10, -4, -3)$.

The corresponding results are listed below:

| Experiments | Dual solution $(\bar{\varsigma}, \bar{\sigma}_1)$ | $\Pi^g(\bar{\varsigma}, \bar{\sigma}_1)$ | $\lambda_{min}(\bar{\varsigma}, \bar{\sigma}_1)$ | Primal solution (\bar{x}, \bar{v}) | $P(\bar{x}, \bar{v})$ |
|-------------|---|--|--|--|-----------------------|
| Example 7 | $\bar{\varsigma} = -0.5$, $\bar{\sigma}_1 = (2.5, 9.75, 6)$. | -33.875 | 1.58694 | $\bar{x} = (1, 1, 1)$, $\bar{v} = (1, 1, 1)$. | -33.875 |
| Example 8 | $\bar{\varsigma} = 0.088$, $\bar{\sigma}_1 = (0, 1.994, 0, 0, 0)$. | -32.8777 | 5.54327 | $\bar{x} = (0.556, 0, 0.978, -0.174, -0.225)$, $\bar{v} = (1, 0, 1, 1, 1)$. | -32.8777 |

Table 7: Results by solving problem (\mathcal{P}_+^g) in Examples 7 and 8.

From Table 7, it holds that $\bar{\varsigma} \geq -\alpha$, $\bar{\sigma}_1 \geq 0$, $G(\bar{\varsigma}, \bar{\sigma}_1) \succ 0$, $f_i + (\bar{\sigma}_1)_i \neq 0$, $\forall i = 1, \dots, n$ in every example, so $(\bar{\varsigma}, \bar{\sigma}_1) \in S_{\varsigma\sigma_1}^+$. By Theorem 4, the solution of primal problem is obtained. So our method is also effective for the general problems.

7 Conclusions and further work

In this paper we propose a canonical duality method for solving a mixed-integer nonconvex fourth-order polynomial minimization problem with fixed cost terms. By rewriting the box constraints in a relaxed quadratic form, a relaxed reformulation of the primal problem (\mathcal{P}_b) is obtained, then the canonical dual problem (\mathcal{P}^\sharp) is defined and the complementary-dual principle is proved. The primal problem (\mathcal{P}_b) is canonically dual to a concave maximization problem (\mathcal{P}_+^\sharp) where a useful space S_+^\sharp is introduced. This constrained nonconvex problem (\mathcal{P}_b) in \mathbb{R}^{2n} can be transformed into a continuous concave maximization dual problem (\mathcal{P}_+^g) in \mathbb{R}^{n+1} without duality gap. The global optimality conditions are proposed and the existence and uniqueness criteria are illustrated. Application to the decoupled mixed-integer problem is illustrated and analytic solution for a global minimum is obtained under some suitable conditions. Several examples are given to show our method is effective. Canonical duality theory is a potentially powerful methodology, which can be used to model complex systems to a wide class of discrete and continuous problems in global optimization and nonconvex analysis. The ideas and results with canonical duality theory presented in this paper can be used or generalized for solving other type of problems in the future.

References

- [1] Padberg M.W., Van Roy T.J., Wolsey L.A.: Valid linear inequalities for fixed charge problems. Oper. Res. 33, 842-861(1985).
- [2] Aardal, K.: Capacitated facility location: separation algorithms and computational experience. Math. Program. 81(2) Ser.B, 149-175(1998).
- [3] Kim D., Pardalos P.M.: A solution approach to the fixed charge network flow problem using a dynamic slope scaling procedure. Operations Research Letters. 24:195-203(1999).

- [4] Glover F., Sherali H.D.: Some classes of valid inequalities and convex hull characterizations for dynamic fixed-charge problems under nested constraints. *Ann. Oper. Res.* 40(1), 215-234(2005).
- [5] Rebennack S., Nahapetyan A., Pardalos P.M.: Bilinear modeling solution approach for fixed charge network flow problems. *Optim Lett.* 3:347-355(2009).
- [6] Vyve M.V.: Fixed-charge transportation on a path: optimization, LP formulations and separation. *Math. Program., Ser. A.* 142:371-395(2013).
- [7] Agra A., Doostmohammadi M.: Facets for the single node fixed-charge network set with a node set-up variable. *Optim Lett.* 8:1501-1515(2014).
- [8] Adlakha V., Kowalski K.: Fractional Polynomial Bounds for the Fixed Charge Problem. *J Optim Theory Appl.* 164:1026-1038(2015).
- [9] Housh M., Cai X.M.: Successive smoothing algorithm for solving large-scale optimization models with fixed cost. *Ann Oper Res.* 229:475-500(2015).
- [10] Gao D.Y., Ruan N., Sherali H.D.: Canonical Dual Solutions for Fixed Cost Quadratic Programs. in: Chinchuluun A., Pardalos P.M., Enkhbat R., Tseveendorj I. (eds.). *Optimization and Optimal Control*, pp. 139-156, Springer, New York(2010).
- [11] Gao, D.Y.: Canonical dual transformation method and generalized triality theory in nonsmooth global optimization. *J. Global Optim.* 17, 127-160(2000).
- [12] Fang S.C., Gao D.Y., Shue R.L., Wu S.Y., Canonical dual approach to solving 0-1 quadratic programming problems. *J. Ind. Manage. Optim.* 4(1), 125-142(2008).
- [13] Gao D.: Canonical duality theory and solutions to constrained nonconvex quadratic programming. *J. Global Optim.* 29, 377-399(2004).
- [14] Gao D., Ruan, N.: On the Solutions to quadratic minimization problems with box and integer constraints. *J. Global Optim.* 47, 463-484(2010).
- [15] Gao D.Y.: Solutions and optimality to box constrained nonconvex minimization problems. *J. Ind. Manage Optim.* 3(2), 293-304(2007).
- [16] Gao D.: Complete solutions and extremality criteria to polynomial optimization problems. *J. Global Optim.* 35, 131-143(2006).
- [17] Zhou X.J., Gao D.Y., Yang C.H.: Canonical primal-dual algorithm for solving fourth-order polynomial minimization problems. *Appl.Math.Comput.* 227:246-255(2014).
- [18] Gao D., Ruan, N., Pardalos, P.: Canonical dual solutions to sum of fourth-order polynomials minimization problems with applications to sensor network localization. In: Boginski, V. L., Commander, C. W., Pardalos, P. M., Ye, Y. (eds.) *Sensors: Theory, Algorithms, and Applications*. Springer, Berlin, pp. 37-54 (2012).
- [19] Wang Z., Fang S., Gao D., Xing W.: Canonical dual approach to solving the maximum cut problem. *J. Global Optim.* 54(2), 341-351(2012).
- [20] Gao, D.Y., Ruan, N., Sherali, H.D.: Solutions and optimality criteria for nonconvex constrained global optimization problems with connections between canonical and Lagrangian duality. *J. Global Optim.* 45:473-497(2009).